

# 陈浩泷

18316035908 · [haolongchen1@link.cuhk.edu.cn](mailto:haolongchen1@link.cuhk.edu.cn) · [github.com/IfReasonable](https://github.com/IfReasonable)

## 教育背景

香港中文大学（深圳），计算机与信息工程，博士	2028.6
• 指导老师：朱光旭, 张纵辉	
暨南大学，软件工程，学士	2023.9
• 指导老师：杨光华, 张鑫源	

## 经历

深圳市大数据研究院, 访问学生	2023.5 – 至今
• 研究方向：大模型高效推理，大模型高效训练，时空数据分析，人工智能在无线通信中的应用	
暨南大学超级计算机团队, 团队成员	2021.6 – 2023.5
• 参加国际高性能计算机比赛 ASC21, SC21.	

## 论文

### An overview of domain-specific foundation model: key technologies, applications and challenges

- Haolong Chen, Hanzhi Chen, Zijian Zhao, Kaifeng Han, Guangxu Zhu, Yichen Zhao, Ying Du, Wei Xu, Qingjiang Shi
- SCIS (CCF A) 已接收。

### STM3: Mixture of Multiscale Mamba for Long-Term Spatio-Temporal Time-Series Prediction

- Haolong Chen, Liang Zhang, Zhengyuan Xin, Guangxu Zhu
- 投稿至 ACM KDD。

### AdaMeZO: Adam-Styled Zeroth-Order Optimizer for LLM Fine-tuning Without Memorizing the Moments

- Zhijie Cai\*, Haolong Chen\*, Guangxu Zhu
- 投稿至 NeurIPS。

### FeedSign: Robust Full-parameter Federated Fine-tuning of Large Models with Extremely Low Communication Overhead of One Bit

- Zhijie Cai\*, Haolong Chen\*, Guangxu Zhu
- 投稿至 IEEE TMC (CCF A)。

### A Semi-Supervised Approach for Telecom Poor User Experience Root-Cause Classification

- Qizhe Li, Haolong Chen, Jiansheng Li, Shuqi Chai, Guangxu Zhu
- 准备投稿期刊。

### First Token Probability Guided RAG for Telecom Question Answering

- Tingwei Chen, Jiayi Chen, Zijian Zhao, Haolong Chen, Liang Zhang, Guangxu Zhu

## 项目

### Multi-Agent 高效协作 | LLM 推理, Multi-Agent 协作

- 从信息多样性角度构建高效 Multi-Agent Debate 系统。

### LLM Routing | LLM 推理, LLM 路由

- 使用经济学拍卖模型建模 LLM Routing 任务负载场景。

### 边缘智能高效推理系统 | LLM 推理, 边缘智能

- 开发了推理增强的边缘推理框架，利用云端部署的大模型的推理输出构建 RAG 语料库，从而无缝增强边缘部署的小模型的能力。

## 边缘智能高效训练系统 | LLM 训练, 边缘智能

- 在边缘智能场景中为大规模 LLM 的训练构建了一个联邦学习系统。针对边缘设备显存低、计算速度慢, 联邦学习系统通信开销大的问题, 构建了一种使用基于种子的随机化来压缩梯度的算法。
- 通过记录少量历史种子的方式, 以极低显存开销实现了二阶动量估计, 大大地加快了训练速度。

## 广东省基础与应用基础研究重大项目：基于环境感知的 6G 网络关键技术研究 | 通信 KPI 预测

- 负责用户时空分布和流量建模与仿真子项目, 该项目是网元时空状态建模与仿真子课题的一部分。
- 开发了一种使用多个基站的联合时空流量建模方法, 并提出了一种新的时空流量预测模型, 该模型集成了 Mamba 长序列神经网络、动态图卷积网络和稀疏混合专家。

## 国家重点研发基金：学习优化理论与方法及其在 5G 网络中的应用 | 通信 KPI 预测

- 负责子项目“5G 网络系统性能建模”中基于时空集成的用户端性能建模。
- 提出了一种基于多模态大语言模型的时空用户性能建模方法, 该方法可以将时间序列用户性能数据与数据集的文本描述和服务区域周围的网络环境信息相结合, 实现高精度预测。

## 多维度用户体验建模（华为 - 深圳市大数据研究院）| 通信 KPI 预测

- 异常类型分类: 通过对通信用户在使用过程中的实时产生的用户测量报告, 构建一个针对用户体验异常的多变量时间序列分类模型。使用这个分类模型, 可以自动地找到用户的异常类型, 帮助进一步做异常根因分析。在此方面, 提出了一种基于阈值标签的条件扩散模型数据增强方法, 从而构造了一个用通信知识增强的半监督训练算法。
- 异常根因分析: 在拥有了分类模型后, 通过 SHAP 方法, 对输入特征进行分析, 找出究竟是什么字段导致了这个体验异常问题。

## 基于 MIMO 信道的频谱效率预测（华为 – 深圳市大数据研究院）| 通信 KPI 预测

- 谱效预测模型: 利用来自多个网格和小区的真实 5G MIMO 测量数据, 构建了一个基于 Transformer 的神经网络模型, 实现从底层 KPI (RSRP, SINR) 到上层 KQI (频谱效率) 的映射, 以多网格和多小区场景下的频谱效率。
- 信道生成模型: 使用信道 – KPI 匹配数据, 构建一个条件扩散模型, 使用 KPI 数据作为控制信号来生成信道数据, 从而填补数据中信道 – KPI 不匹配的部分。

## 国际学术会议审稿

- NeurIPS 25, ICASSP 24, ICC 24,25, GLOBECOM 25, ICCC 25, WCNC 24,25, PIMRC 25 等。

## 专利

### 无线通信中的流量预测方法、装置、电子设备及存储介质

- 陈浩泷, 辛政沅, 朱光旭, 深圳市大数据研究院, 专利号: ZL202511087894.7, 授权日: 2025 年 10 月 28 日。

### 基于多模态大模型的通信指标预测方法及相关设备

- 陈浩泷, 朱光旭, 史清江, 深圳市大数据研究院, 专利号: ZL202510542918.7, 授权日: 2025 年 10 月 17 日。

### 模型训练方法、文本分类方法、装置、电子设备及介质

- 蔡智捷, 陈浩泷, 朱光旭, 张纵辉, 史清江, 深圳市大数据研究院, 专利号: ZL2025103509755, 授权日: 2025 年 9 月 16 日。

### 通信和内存高效的大模型分布式训练方法、文本分类方法

- 蔡智捷, 陈浩泷, 朱光旭, 深圳市大数据研究院, 专利申请号: ZL2025100670425, 授权日: 2025 年 8 月 12 日。

### 频谱效率的预测方法、装置、电子设备及存储介质

- 陈浩泷, 朱光旭, 史清江, 深圳市大数据研究院, 专利号: ZL2023115716969, 授权日: 2024 年 2 月 23 日。

## 软件专著

### 基于大规模实测数据源的谱效预测算法半监督训练及求解系统 V1.0

- 深圳市大数据研究院, 登记号: 2024SR1450315, 授权日: 2024 年 9 月 29 日。

## 技能

---

### 编程

- 熟练: Python, PyTorch, Linux
- 熟悉: Matlab, C/C++, MySQL, Git, Java, 网页前端开发, 网页后端开发, TensorFlow

### 语言

- 英语 (雅思: 6.5, CET-4: 548, CET-6: 542)
- 中文 (母语)